Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Linear Algebra and its Applications 413 (2006) 303–318

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Simultaneous versal deformations of endomorphisms and invariant subspaces ☆

Albert Compta [a], Uwe Helmke [b], Marta Peña [a], Xavier Puerta [a,*]

[a] *Departament de Matemàtica Aplicada I, ETS Enginyeria Industrial de Barcelona, UPC Diagonal 647, 08028 Barcelona, Spain*

[b] *Department of Mathematics, University of Würzburg, Am Hubland, 97074 Würzburg, Germany*

Received 2 November 2004; accepted 5 June 2005

Available online 22 September 2005

Submitted by I. Zaballa

Abstract

We study the set $\mathscr{M}$ of pairs (f, V), defined by an endomorphism f of $\mathbb{F}^n$ and a d-dimensional f-invariant subspace V. It is shown that this set is a smooth manifold that defines a vector bundle on the Grassmann manifold. We apply this study to derive conditions for the Lipschitz stability of invariant subspaces and determine versal deformations of the elements of $\mathscr{M}$ with respect to a natural equivalence relation introduced on it.

Keywords: Grassmann manifold; Lie group action; Lipschitz stability; Versal deformation

1. Introduction

The set of d-dimensional invariant subspaces of a given linear endomorphism of a finite dimensional vector space has been extensively studied by many authors. The starting point to our investigation is [8], where an explicit stratification of the set into manifolds is constructed, that are defined by fixing the Segre characteristic of

☆ Partially supported by DGICYT n.PB97-0599-C03-03, as well as DAAD D/0243869.

* Corresponding author.

E-mail address: francisco.javier.puerta@upc.es (X. Puerta).

0024-3795/$ - see front matter © 2005 Published by Elsevier Inc.

doi:10.1016/j.laa.2005.06.017

the induced restriction of the linear operator to the subspace. The singularities of the union of these strata constitute an obstruction for the implementation of Arnold's techniques in the study of local perturbations of invariant subspaces. Partial results in this direction are obtained by [5].

Surprisingly, and in contrast to the presence of singularities in the set of invariant subspaces, the set $\mathscr{M}$ of pairs, formed by an endomorphism together with an invariant subspace of fixed dimension, turns out to be a smooth manifold. Moreover, natural projections both into the set of invariant subspaces of a fixed endomorphism as well as into the set of endomorphisms that leave a given subspace invariant, enable us to study these two situations simultaneously. Therefore, the study of the local perturbations of the above pairs seems to be a natural approach to the study of local perturbations in the image of the above projections. In particular, it is possible to obtain simple sufficient conditions for Lipschitz-stability of an invariant subspace.

In general, given a Lie group action on a smooth manifold M, versal and miniversal deformations with regard to this action were first introduced by Arnold in [1] (see also [2]), and have been subsequently studied in several areas of linear algebra related to local perturbation analysis. In particular, Arnold's technique for studying the versal deformations of a square matrix with regard to the similarity group action, has been generalized to (A, B) pairs with regard to feedback equivalence [4], and to square matrices having a fixed zero block structure, with regard to the restricted similarity equivalence [3]. This last set of matrices corresponds with the set of endomorphisms that keep invariant a fixed subspace. In this paper, we compute a miniversal deformation of endomorphism/invariant subspace pairs with regard to a natural equivalent relation defined on it. This allows us to construct in a rather straightforward way both versal deformations of square matrices with a fixed zero block structure as well as versal deformations of invariant subspaces for a fixed endomorphism.

The paper is organized as follows. In Section 2 we prove that the set $\mathscr{M}$ of pairs formed by an endomorphism and an invariant subspace of it of a fixed dimension, is a smooth manifold. We compute the tangent spaces and construct explicit local coordinate charts of $\mathscr{M}$. As an application, we derive in Section 3 sufficient conditions for the Lipschitz-stability of an invariant subspace. In Section 4 we construct miniversal deformations of elements in $\mathscr{M}$ with regard to a natural Lie group action. In Section 5 we derive from the above deformation a miniversal deformation of endomorphisms with a given invariant subspace. We use the following notation. $\mathbb{F}$ denotes both the sets of real and complex numbers, respectively. $M_{n,m}$ denotes the set of $n \times m$ matrices with entries in $\mathbb{F}$, and $M^*_{n,m}$ denotes the set of full rank ones. The set of square $n \times n$ matrices is denoted by M_n. We denote the general linear group of $n \times n$ matrices by Gl_n. $\mathrm{Gr}_d(X)$ is the Grassmann manifold formed by the set of all d-dimensional subspaces of X. If V is a subspace of X, we say that a basis of X is adapted to V, if it is obtained by extending a basis of V.

2. The manifold of pairs endomorphism–invariant subspace

We consider the set of pairs (f, V) where f is an element of M_n and V is a d-dimensional invariant subspace of f. We denote this set by $\mathscr{M}$. It is clear that $\mathscr{M} \subset M_n \times \mathrm{Gr}_d(\mathbb{F}^n)$. We denote the last product by $\mathscr{N}$. In this section we prove that $\mathscr{M}$ is a smooth submanifold of $\mathscr{N}$ of dimension n^2, and we give a local parameterization of it. The first step consists in identifying the Grassmann manifold $\mathrm{Gr}_d(\mathbb{F}^n)$ with the following set of selfadjoint projection operators

$$\mathscr{P}_d = \{P \in M_n(\mathbb{F}) | P^* = P, P^2 = P, \mathrm{rank}\, P = d\}.$$

Let $M_n^d = \{A \in M_n \,|\mathrm{rank}\, A = d\}$. We make use of the following preliminary facts. For a proof we refer to [7].

Proposition 2.1. *With the above notation, we have that*

(i) M_n^d *is a smooth submanifold of* M_n *of dimension* $n^2 - (n-d)^2$.
(ii) $\mathscr{P}_d$ *is a smooth submanifold of* M_n^d *of dimension* $d(n-d)$.
(iii) $T_P\mathscr{P}_d = \{[P, \Omega] \,|\Omega = -\Omega^*, \Omega \in M_n\}$.
(iv) $\mathscr{P}_d$ *and* $\mathrm{Gr}_d(\mathbb{F}^n)$ *are diffeomorphic manifolds.*

Thanks to the above proposition, we identify, from now on, the Grassmann manifold $\mathrm{Gr}_d(\mathbb{F}^n)$ with $\mathscr{P}_d$. Since the condition $f(\mathrm{Im}\, P) \subset \mathrm{Im}\, P$ is equivalent to $(I - P)fP = 0$,

$$\mathscr{M} = \{(f, P) \in \mathscr{N} \,|(I - P)fP = 0\}.$$

The main result of this section can now be stated as follows.

Theorem 2.2. *With the above notation*, $\mathscr{M}$ *is a smooth*, *closed submanifold of* $\mathscr{N}$ *of dimension* n^2 *and the tangent space to an element* $(f, P) \in \mathscr{M}$ *is*

$$\begin{aligned} T_{(f,P)}\mathscr{M} = \{(g, [P, \Omega])| g, \Omega \in M_n(\mathbb{F}), \Omega = -\Omega^*, (I - P)gP \\ -[P, \Omega]fP + (I - P)f[P, \Omega] = 0\}. \end{aligned}$$

Proof. Note, that $\mathscr{M}$ is described as the solution set to a system of real polynomial equations in $\mathscr{N}$ and therefore is a closed subset of $\mathscr{N}$. We begin by constructing local coordinate charts for $\mathscr{M}$.

Given $(f_0, V_0) \in \mathscr{N}$, we consider a basis of $\mathbb{F}^n$ such that $V_0 = \mathrm{Im} \begin{pmatrix} I_d \\ 0 \end{pmatrix}$. We identify each endomorphism of $\mathbb{F}^n$ with its matrix with regard to this basis. If $f = \begin{pmatrix} A & C \\ D & B \end{pmatrix}$ with $A \in M_d$, $C \in M_{d,n-d}$, $D \in M_{n-d,d}$ and $B \in M_{n-d}$, we denote the blocks A, C, D, B by $f^{(1)}, f^{(2)}, f^{(3)}$ and $f^{(4)}$, respectively. Let $\mathscr{N}_0 = \{(f, V) \,| f \in$

$M_n, V = \operatorname{Im}\begin{pmatrix} I_d \\ Q \end{pmatrix}, Q \in M_{n-d,d}\}$. It is clear that $\mathscr{N}_0$ is an open set of $\mathscr{N}$ containing (f_0, V_0).

Obviously, the map $\gamma : \mathbb{F}^{n^2+d(n-d)} \to \mathscr{N}_0$ defined by

$$\gamma(A, B, C, D, Q) = \left(\begin{pmatrix} A & C \\ D & B \end{pmatrix}, \operatorname{Im}\begin{pmatrix} I_d \\ Q \end{pmatrix}\right)$$

is a local coordinate system of $\mathscr{N}$ with image $\mathscr{N}_0$. We now show that it also defines a coordinate chart for $\mathscr{M}$. Let $\mathscr{M}_0 := \mathscr{M} \cap \mathscr{N}_0$. From the above the following holds true. $\square$

Lemma 2.3. *With the above notation,*

$$\mathscr{M}_0 = \left\{\left(\begin{pmatrix} A & C \\ D & B \end{pmatrix}, \operatorname{Im}\begin{pmatrix} I_d \\ Q \end{pmatrix}\right) \text{ with } D = QA - BQ + QCQ\right\}.$$

Let $\psi : \mathbb{F}^{n^2} \to \mathbb{F}^{n^2+d(n-d)}$ be the map defined by

$$\psi(A, B, C, Q) = (A, B, C, QA - BQ + QCQ, Q).$$

We denote $\gamma \circ \psi$ by θ and we have the following proposition.

Proposition 2.4. *θ is a local coordinate system of $\mathscr{M}$ with $\theta(\mathbb{F}^{n^2}) = \mathscr{M}_0$.*

Proof. From Lemma 2.3, θ is injective and its image is $\mathscr{M}_0$.

It is clear that θ is differentiable. In order to prove that θ is a diffeomorphism it is sufficient to prove that $\mathrm{d}\psi_{(A,B,C,Q)}$ is injective. This follows from the equality

$$\begin{aligned}\mathrm{d}\psi_{(A,B,C,Q)}(\dot{A}, \dot{B}, \dot{C}, \dot{Q}) = (\dot{A}, \dot{B}, \dot{C}, &(QC - B)\dot{Q} + \dot{Q}(A + CQ) + Q\dot{A} \\ &+ Q\dot{C}Q - B\dot{Q}, \dot{Q}). \qquad \square\end{aligned}$$

This completes the construction of the coordinate charts. It is easily seen, although a bit tedious to show, that these coordinates charts glue together to define an atlas on $\mathscr{M}$. We therefore proceed to verify the formula for the tangent spaces. $\mathscr{M}$ is the inverse image of zero by the smooth map

$$\begin{aligned} \varphi : \mathscr{N} &\to M_n \\ (f, P) &\longmapsto (I - P)fP. \end{aligned}$$

Now we prove that $\mathrm{d}\varphi_{(f,P)}$ has constant rank. For this, we evaluate

$$\mathrm{d}\varphi_{(f,P)}(\dot{f}, \dot{P}) = (I - P)\dot{f}P - \dot{P}fP + (I - P)f\dot{P}.$$

Since $\dot{P} = [P, \Omega]$ and $\Omega = -\Omega^*$, we have that

$$\begin{aligned}\operatorname{Im} \mathrm{d}\varphi_{(f,P)} = \{&(I - P)\dot{f}P - [P, \Omega]fP + (I - P)f[P, \Omega] \mid \dot{f}, \\ &\Omega \in M_n, \Omega = -\Omega^*\} \subset M_n.\end{aligned}$$

In order to see that rank $\mathrm{d}\varphi_{(f,P)}$ is constant we compute the dimension of the orthogonal of its image in M_n. One has that $L \in (\mathrm{Im}\,\mathrm{d}\varphi_{(f,P)})^{\perp}$ if and only if

$$\operatorname{trace}(L^*(I-P)\dot{f}P - L^*[P,\Omega]fP + L^*(I-P)f[P,\Omega]) = 0$$
$$\text{for all } \dot{f}, \Omega \in M_n(\mathbb{F}),\ \Omega = -\Omega^*.$$

This is equivalent to the conditions

$$\operatorname{trace}(PL^*(I-P)\dot{f}) = 0 \quad \text{for all } \dot{f} \in M_n$$

and

$$\operatorname{trace}(L^*(I-P)f[P,\Omega] - L^*[P,\Omega]fP) = 0 \quad \text{for all } \Omega \in M_n, \Omega = -\Omega^*.$$

These conditions in turn are equivalent to

(1) $PL^*(I-P) = 0$,
(2) $\operatorname{trace}(-L^*(I-P)f\Omega P - L^*P\Omega fP + L^*\Omega fP) = 0$.

Notice that $\operatorname{trace}(-PL^*(I-P)f\Omega + \Omega fPL^*(I-P)) = 0$, and therefore (1) implies (2). Therefore,

$$(\mathrm{Im}\,\mathrm{d}\varphi_{(f,P)})^{\perp} = \{L \in M_n(\mathbb{F}) \mid (I-P)LP = 0\}.$$

In order to compute the dimension of $\mathrm{Im}\,\mathrm{d}\varphi_{(f,P)}$, we consider a basis of $\mathbb{F}^n$, so that $P = \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix}$, then decomposing L according to the blocks of P, the equation $(I-P)LP = 0$ is

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & I_{n-d} \end{pmatrix} \begin{pmatrix} L_1 & L_2 \\ L_3 & L_4 \end{pmatrix} \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ L_3 & 0 \end{pmatrix},$$

which is equivalent to $L_3 = 0$.

This implies that rank $\mathrm{d}\varphi_{(f,P)} = d(n-d)$ and, therefore, since $\mathscr{M}$ is a smooth manifold of dimension n^2, we conclude that

$$\dim \mathscr{M} = \dim \mathscr{N} - \operatorname{rank} \mathrm{d}\varphi_{(f,P)} = (n^2 + nd - d^2) - d(n-d) = n^2.$$

Using the rank formula we conclude that dim ker $\mathrm{d}\varphi_{(f,P)} = \dim T_{\mathscr{M}}$ and therefore the tangent space formula holds.

3. Stability of invariant subspaces

We keep the notation of the previous section. Consider the following diagram

$$\begin{array}{ccc} \mathscr{M} & \xrightarrow{\pi_1} & M_k \\ \pi_2 \downarrow & & \\ \mathrm{Gr}_d(\mathbb{F}^n) & & \end{array}$$

where π_1 and π_2 are the natural projection operators. In the sequel we will study an amplification of the concept of stable invariant subspaces, introduced by [6]. Let $\Theta(V, V')$ denote the *gap distance* between two linear subspaces.

Definition 3.1. Let $(f, V) \in \mathscr{M}$. Then V is called stable, if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\|f' - f\| < \delta$ implies that exists an f'-invariant subspace V' with $\Theta(V', V) < \varepsilon$. The subspace V is called locally Lipschitz-continuous, if, locally around (f, V), there exists $L > 0$ such that $\Theta(V', V) \leqslant L\|f' - f\|$.

Obviously, the concept of Lipschitz stability is stronger than the purely topological notion of stability. In [6], stable invariant subspaces are extensively studied. In particular, necessary and sufficient conditions that characterize stable invariant subspaces are derived. In this section we show how through an analysis of the smooth projection maps π_1 and π_2 one can derive a simple sufficient condition also for the Lipschitz-stability of an invariant subspace.

Proposition 3.2. *Let $(f, V) \in \mathscr{M}$. If $\mathrm{d}\pi_{1,(f,V)}$ is bijective, then V is Lipschitz-stable.*

Proof. Let us denote $\mathrm{d}\pi_{1,(f,V)}$ by α. If α is bijective, from the inverse function theorem, there exist open sets $\mathscr{U}$ and $\mathscr{V}$ with $(f, V) \in \mathscr{U} \subset \mathscr{M}$ and $f \in \mathscr{V} \subset M_n(\mathbb{F})$ such that $\alpha : \mathscr{U} \to \mathscr{V}$ is a diffeomorphism. We consider the composition

$$\begin{array}{ccccc} \mathscr{V} & \stackrel{\alpha^{-1}}{\to} & \mathscr{U} & \stackrel{\pi_2}{\to} & \mathrm{Gr}_d(\mathbb{F}^n) \\ f' & \longmapsto & (f', V') & \longmapsto & V' \end{array}$$

Taking into account that α^{-1} and π_2 are smooth and hence Lipschitz-continuous maps, the result follows. $\square$

Next we prove necessary and sufficient conditions for the bijectivity of $\mathrm{d}\pi_{1,(f,V)}$ which, thanks to the above proposition, yield sufficient conditions for the Lipschitz-stability of V.

Proposition 3.3. *Let $(f, V) \in \mathscr{M}$, and P be the selfadjoint projection operator representing the subspace V. Let $\begin{pmatrix} A_0 & C_0 \\ 0 & B_0 \end{pmatrix}$ denote the matrix of f in any basis adapted to V, and $\Omega \in M_n$. Then, the following statements are equivalent:*

(i) $\mathrm{d}\pi_{1,(f,V)}$ *is bijective.*
(ii) A_0 *and* B_0 *have disjoint sets of eigenvalues.*
(iii) $-P\Omega f P + \Omega f P - f\Omega P + P f\Omega P = 0$, *with* $\Omega = -\Omega^*$ *implies* $P\Omega = \Omega P$.

Proof. In a basis adapted to V, we may assume that $V = \mathrm{Im}\begin{pmatrix} I_d \\ 0 \end{pmatrix}$, so that the coordinates of (f, V) in the local coordinate system θ given in Proposition 2.4 are (A, B, C, Q). It is easily checked that in this coordinate system, the map π_1, denoted by $\dot{\pi}_1$, is given by

$$\dot{\pi}_1(A, B, C, Q) = (A, B, C, QA - BQ + QCQ).$$

Since $\mathrm{d}\dot{\pi}_{1(A_0,B_0,C_0,0)}(\dot{A}, \dot{B}, \dot{C}, \dot{Q}) = (\dot{A}, \dot{B}, \dot{C}, \dot{Q}A_0 - B_0\dot{Q})$, it is clear that $\mathrm{d}\dot{\pi}_{1(A_0,B_0,C_0,0)}$ is bijective if and only if the map $D \mapsto DA_0 - B_0D$ is bijective. By a well-known result about injectivity of Sylvester operators this is equivalent to A_0 and B_0 having disjoint spectrum. This completes the proof of the equivalence between (i) and (ii). From the description of $\mathscr{M}$ and $T_{(f,P)}\mathscr{M}$ given in the previous section, we conclude that $\mathrm{d}\pi_{(f,P)}$ is bijective if and only if $-[P, \Omega]fP + (I - P)f[P, \Omega] = 0$ implies $[P, \Omega] = 0$. This shows the equivalence between (i) and (ii). The result follows. $\square$

The above result characterizes the critical points of the projection map onto the first factor. In contrast, the projection map onto the second factor has better geometric properties, as it has no critical points.

Proposition 3.4. *With the above notation, for every* $(f, V) \in \mathscr{M}$, $\mathrm{rank}\, \mathrm{d}\pi_{2,(f,V)} = d(n - d)$.

Proof. As in the proof of Proposition 3.3 we can assume that the coordinates of (f, V) are (A, B, C, Q) and that the projection operator π_2 is given in its local coordinate system of $\mathscr{M}$ and the corresponding coordinate system of $\mathrm{Gr}_d(\mathbb{F}^n)$, by $\dot{\pi}_2(A, B, C, Q) = Q$. Hence, $\mathrm{d}\dot{\pi}_2(\dot{A}, \dot{B}, \dot{C}_0, \dot{Q}) = \dot{Q}$ so that $\mathrm{rank}\, \mathrm{d}\pi_{2,(f,V)} = d(n - d)$. $\square$

Corollary 3.5. *The map* π_2 *induces a submersion* $\mathscr{M} \to \mathrm{Gr}_d(\mathbb{F}^n)$. *Moreover*, π_2 *defines a smooth vector bundle on the Grassmannian.*

From this corollary, we derive the following proposition. Roughly speaking, it states that any map f is ‘stable’ with regard to an invariant subspace V of it.

Proposition 3.6. *Given* $(f, V) \in \mathscr{M}$, *for every* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that for any* $V' \in \mathrm{Gr}_d(\mathbb{F}^n)$ *with* $\Theta(V, V') < \delta$, *there exists* f' *with* $(f', V') \in \mathscr{M}$ *and* $\|f' - f\| < \varepsilon$.

Proof. Let $\sigma : V \to \mathscr{M}$ be a local section of π_2 in a neighbourhood of V. Then, the proposition follows from the continuity of the map $\pi_1 \circ \sigma$. $\square$

4. Miniversal deformations

Given a smooth manifold M and a Lie group G acting on it, versal and miniversal deformations of an element $x \in M$ with regard to this action, are concepts introduced by Arnold in order to study local perturbations of x.

Definition 4.1. A *versal deformation* of x with regard to the action defined by G is a smooth map $\varphi : \mathscr{U} \to M$ where $\mathscr{U}$ is an open neighborhood of 0 in $\mathbb{R}^l$ and $\varphi(0) = x$, such that for any map $\psi : \mathscr{V} \to M$ of the same kind, there exist a smooth map $\alpha : \mathscr{V}' \to \mathscr{U}$ with $\mathscr{V}' \subset \mathscr{V}, 0 \in \mathscr{V}'$, $\alpha(0) = 0$ and a map $h : \mathscr{V}' \to G$ with $h(0) = \mathrm{i}d$, such that $\psi(t) = h(t) * \varphi(h(t))$ (the action of G on M is denoted by $*$). If l is minimal, φ is called *miniversal.*

We refer the reader to [1] for preliminaries relevant for our purposes. In this section, we make use of the following result. We denote by $\mathscr{O}(x)$ the orbit of x with regard the considered action.

Theorem 4.2 [1]. *A versal deformation of $x \in M$ is a parameterization of any manifold N transversal to $\mathscr{O}(x)$ at x, that is to say, a manifold N such that*

$$T_x(N) + T_x\mathscr{O}(x) = T_x(M).$$

Furthermore, this versal deformation is miniversal if and only if the above sum is direct. In this case $\dim \mathscr{O}(x) = \dim M - \dim N$.

We consider here the action of Gl_n on $\mathscr{N}$ defined by

$$\sigma(S, (f, V)) = (SfS^{-1}, S(V)).$$

Notice that $f(V) \subset V$ implies $(SfS^{-1})(S(V)) \subset S(V)$ and therefore, σ can be restricted to an action on $\mathscr{M}$ denoted also by σ. We denote the orbit $\sigma(\mathrm{Gl}_n, (f, V))$ by $\mathscr{O}(f, V)$.

In order to compute a miniversal deformation of an element of $\mathscr{M}$ with regard to the above group action, we make use of the local description of $\mathscr{M}$ (or $\mathscr{N}$) given in Section 2. Following the notation introduced there, we first compute the coordinates of the tangent space of $\mathscr{O}(f_0, V_0)$ in (f_0, V_0). Since the map $(f, V) \mapsto (SfS^{-1}, S(V))$ is a diffeomorphism for all $S \in \mathrm{Gl}_n$, we can compute, with out loss of generality, the miniversal deformation in a particular element of the orbit of $(f_0, V_0) \in \mathscr{M}$. So, for simplicity, we take $V_0 = \mathrm{Im}\begin{pmatrix} I \\ 0 \end{pmatrix}$ and $f_0 = \begin{pmatrix} A_0 & C_0 \\ 0 & B_0 \end{pmatrix}$. Notice that A_0, B_0, C_0 are not uniquely defined. In fact, although taking a suitable S we can reduce A_0 and B_0 to Jordan matrices, there is not known a canonical form for C_0. For this reason, we make not any assumption on the particular forms of the matrices A_0, B_0, C_0. We prove the following proposition.

Proposition 4.3. *With the above notation,* $T_{(A_0,B_0,C_0,0)}(\theta^{-1}(\mathscr{O}(f_0, V_0)) \cap \mathscr{M}_0))$ *is the set*

$$\{(\dot{S}_1 A_0 - A_0\dot{S}_1 - C_0\dot{S}_3, \dot{S}_3 C_0 + \dot{S}_4 B_0 - B_0\dot{S}_4, \dot{S}_1 C_0 + \dot{S}_2 B_0 \\ -A_0\dot{S}_2 - C_0\dot{S}_4, \dot{S}_3)\},$$

where $\dot{S}_1 \in M_d$, $\dot{S}_2 \in M_{d,n-d}$, $\dot{S}_3 \in M_{n-d,d}$, $\dot{S}_4 \in M_{n-d}$.

Proof. For continuity, the action σ restricts to a map, denoted also by σ, $\sigma : \mathscr{G}_0 \times \mathscr{N}'_0 \to \mathscr{N}_0$ where $\mathscr{G}_0$ and $\mathscr{N}'_0$ are open neighborhoods of I_n in Gl_n and of (f_0, V_0) in $\mathscr{N}$, respectively. Analogously, σ defines a map, denoted also by σ, $\sigma : \mathscr{G}_0 \times \mathscr{M}'_0 \to \mathscr{M}_0$, where $\mathscr{M}'_0 = \mathscr{N}'_0 \cap \mathscr{M}$.

Notice that if $\begin{pmatrix} S_1 & S_2 \\ S_3 & S_4 \end{pmatrix} \in \mathscr{G}_0$, $\begin{pmatrix} S_1 & S_2 \\ S_3 & S_4 \end{pmatrix}\begin{pmatrix} I_d \\ 0 \end{pmatrix} \in \mathscr{N}_0$ implies that S_1 is invertible.

Denoting $\alpha = \theta^{-1} \circ \sigma \circ (I_d, \theta)$, we have that $\alpha(\mathscr{G}_0, (A_0, B_0, C_0, 0))$ is an open neighborhood of $(A_0, B_0, C_0, 0)$ in $\theta^{-1}(\mathscr{O}(f_0, V_0)) \cap \mathscr{M}_0)$.

Moreover, if $S = \begin{pmatrix} S_1 & S_2 \\ S_3 & S_4 \end{pmatrix} \in \mathscr{G}_0$, we have that

$$\alpha(S, (A_0, B_0, C_0, 0)) = ((Sf_0S^{-1})^{(1)}, (Sf_0S^{-1})^{(4)}, (Sf_0S^{-1})^{(2)}, S_3S_1^{-1}).$$

Computing the image of the differential of the map $S \mapsto \alpha(S, (A_0, B_0, C_0, 0))$, the proposition follows. $\square$

The main result of this section is given by the following theorem.

Theorem 4.4. *With the above notation, a miniversal deformation of* (f_0, V_0) *is given by the set of pairs*

$$\left(\begin{pmatrix} A_0 + X & C_0 + Z \\ 0 & B_0 + Y \end{pmatrix}, \mathrm{Im}\begin{pmatrix} I_d \\ 0 \end{pmatrix}\right),$$

where X, Y, Z *verify the conditions*

(1) $A_0^* Z - Z B_0^* = 0$,
(2) $[A_0^*, X] - Z C_0^* = 0$,
(3) $[Y, B_0^*] - C_0^* Z = 0$.

Proof. For simplicity, in this proof we denote A_0, B_0 and C_0 by A, B and C, respectively. Let $\mathscr{S}$ be the subspace of $\mathbb{F}^{n^2}$ formed by the elements $(X, Y, Z, 0)$ verifying conditions (1)–(3). We claim that $\mathscr{S}$ is a supplementary subspace of $T_{(A,B,C,0)}(\theta^{-1}(\mathscr{O}(f_0, V_0)) \cap \mathscr{M}_0))$. In order to prove this, we make use of the following lemma.

Lemma 4.5. $T_{(A,B,C,0)}(\theta^{-1}(\mathscr{O}(f_0, V_0)) \cap \mathscr{M}_0))^{\perp}$ *is the set of matrices* (X, Y, Z, T) *verifying conditions* (1)–(3) *and* $T = C^*X - YC^*$.

Proof. $(X, Y, Z, T) \in T_{(A,B,C,0)}(\theta^{-1}(\mathscr{O}(f_0, V_0)) \cap \mathscr{M}_0))^{\perp}$ if and only if

$$\operatorname{trace}(X^*(S_1A - AS_1 - CS_3) + Y^*(S_3C - S_4B - BS_4) \\ + Z^*(S_1C - S_2B - AS_2 - CS_4) + T^*S_3) = 0,$$

for all $\begin{pmatrix} S_1 & S_2 \\ S_3 & S_4 \end{pmatrix} \in M_n$.

Since the trace is invariant by circular permutations of the matrices, the above condition is equivalent to

$$\operatorname{trace}(AX^*S_1 - X^*AS_1 - X^*CS_3 + CY^*S_3 + BY^*S_4 - Y^*BS_4 \\ + Z^*S_1 + BZ^*S_2 - Z^*AS_2 - Z^*CS_4 + T^*S_3) = 0,$$

or, equivalently, $\operatorname{trace}((AX^* - X^*A - CZ^*)S_1 + (BZ^* - Z^*A)S_2) + (-X^*C - CY^* + T)S_3 + (BY^* - Y^*B - Z^*C)S_4) = 0$, which is equivalent to the conditions given in this lemma. □

Thanks to the above lemma, the map $(X, Y, Z, 0) \mapsto (X, Y, Z, C^*X - YC^*)$ defines an isomorphism between $\mathscr{S}$ and $T_{(A,B,C,0)}(\theta^{-1}(\mathscr{O}(f_0, V_0)) \cap \mathscr{M}_0))^{\perp}$. Moreover, since

$$\langle (X, Y, Z, 0), (X, Y, Z, C^*X - YC^*)\rangle = \operatorname{trace}(X^*X + Y^*Y + Z^*Z),$$

we have that $\mathscr{S} \cap T_{(A,B,C,0)}(\theta^{-1}(\mathscr{O}(f_0, V_0)) \cap \mathscr{M}_0)) = \{0\}$ and $\mathscr{S}$ is as we have claimed.

Then, since θ is a diffeomorphism, $\theta(\mathscr{S})$ is a submanifold transversal to $T_{(f_0,V_0)}$ $\mathscr{O}_{(f_0,V_0)}$ and, applying Theorem 4.2 the theorem follows. □

In the manifold $\mathscr{N}$, we can reproduce the same reasoning than for $\mathscr{M}$ with γ instead of θ and with $\mathbb{F}^{n^2+d(n-d)}$ instead of $\mathbb{F}^{n^2}$. Then, it follows a similar result.

Theorem 4.6. *A miniversal deformation of* $(f_0, V_0) \in \mathscr{N}$ *is given by the set of pairs*

$$\left(\begin{pmatrix} A_0 + X & C_0 + Z \\ V & B_0 + Y \end{pmatrix}, \operatorname{Im}\begin{pmatrix} I_d \\ 0 \end{pmatrix}\right),$$

where X, Y, Z *verify conditions* (1)–(3) *given in Theorem* 4.4.

We now apply the above miniversal deformations in order to compute the dimension of the orbit of a pair (f, V). We remark that the codimension of an orbit in $\mathscr{M}$ or in $\mathscr{N}$ differ in $d(n-d)$ (the number of parameters of V). We prove the following proposition.

Proposition 4.7. *Let* M *be the matrix*

$$\begin{pmatrix} I_d \otimes A^t - A \otimes I_d & 0 & -\overline{C} \otimes I_d \\ 0 & I_{n-d} \otimes B^t - B \otimes I_{n-d} & I_{n-d} \otimes C^* \\ 0 & 0 & I_{n-d} \otimes A^t - B \otimes I_d \end{pmatrix}.$$

Then,

$$\dim \mathscr{O}(f, V) = \operatorname{rank} M + d(n-d).$$

Proof. We recall that the vec-operator of a matrix space is the isomorphism

$$\begin{array}{cccc} \mathrm{vec}: & M_{p,q} & \to & M_{pq,1} \\ & X & \longmapsto & (x_{11}, \ldots, x_{1q}, \ldots, x_{p1}, \ldots, x_{pq})^t. \end{array}$$

The proposition follows by checking that conditions (1)–(3) in Theorem 4.4 are equivalent to $\begin{pmatrix} \mathrm{vec}(X) \\ \mathrm{vec}(Y) \\ \mathrm{vec}(Z) \end{pmatrix} \in \ker M$. $\square$

Finally, we obtain lower and upper bounds for the dimension of the orbit of a pair $(f, V) \in \mathscr{M}$ depending on the Segre characteristics of the restriction and the quotient. We can reduce the problem to the case where f has only one eigenvalue. Thus, we can assume that f is nilpotent. We prove the following theorem.

Theorem 4.8. *Let $\mathscr{M}_{\gamma,\beta}$ be the set of pairs $(f, V) \in \mathscr{M}$ with f nilpotent and γ, β the Segre characteristics of the restriction and the quotient induced map defined by f in V and $\mathbb{F}^n/V$, respectively. Then*,

(i) $\max_{(f,V)\in M_{\gamma,\beta}}\{\operatorname{codim} \mathscr{O}(f, V)\} = \sum_{1\leqslant i,j\leqslant r} \min(\gamma_i, \gamma_j) + \sum_{1\leqslant i,j\leqslant s} \min(\beta_i, \beta_j) + \sum_{1\leqslant i\leqslant r, 1\leqslant j\leqslant s} \min(\gamma_i, \beta_j)$.

(ii) $\min_{(f,V)\in M_{\gamma,\beta}}\{\operatorname{codim} \mathscr{O}(f, V)\} = \sum_{1\leqslant i,j\leqslant r} \min(\gamma_i, \gamma_j) + \sum_{1\leqslant i,j\leqslant s} \min(\beta_i, \beta_j)$.

Proof. We make use of the following lemmas.

Lemma 4.9. *Let* $N = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{pmatrix} \in M_\gamma$ *and* $D = \begin{pmatrix} N^t & & & \\ -I & N^t & & \\ & \ddots & \ddots & \\ & & -I & N^t \end{pmatrix} \in M_{\gamma\beta}$. *Then*,

$$\operatorname{rank} D = \gamma\beta - \min(\gamma, \beta).$$

Proof. Let $n(1)$ be the number of elements of the matrix D equal to 1 and $n(-1)$ the number of elements of this matrix equal to -1. Then, we have

$$n(1) = (\gamma - 1)\beta = \gamma\beta - \beta,$$
$$n(-1) = \gamma(\beta - 1) = \gamma\beta - \gamma.$$

Hence, $\operatorname{rank} D \geqslant \max(n(1), n(-1)) = \gamma\beta - \min(\gamma, \beta)$.

We distinguish the cases:

(a) $\beta \geqslant \gamma$
rank $D = \gamma(\beta - 1) = \gamma\beta - \gamma = \gamma\beta - \min(\gamma, \beta)$,
which can be proved by transformations of rows that make the γ first rows zero.
(b) $\beta \leqslant \gamma$
rank $D = \beta(\gamma - 1) = \gamma\beta - \beta = \gamma\beta - \min(\gamma, \beta)$,
which can be proved by transformations of rows and columns that make the matrices $-I$ zero. $\square$

Lemma 4.10. *Let $A \in M_l$ and $B \in M_q$ be Jordan nilpotent matrices, with $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_r)$ the Segre characteristic of A and $\beta = (\beta_1, \dots, \beta_s)$ the Segre characteristic of B. Then,*

$$\operatorname{rank}(I_q \otimes A^t - B \otimes I_l) = lq - \sum_{1 \leqslant i \leqslant r, 1 \leqslant j \leqslant s} \min(\gamma_i, \beta_j).$$

Proof. Since $l = \sum_{1 \leqslant i \leqslant r} \gamma_i$ and $q = \sum_{1 \leqslant j \leqslant s} \beta_j$, the formula that we have to prove is equivalent to

$$\operatorname{rank}(I_q \otimes A^t - B \otimes I_l) = \sum_{1 \leqslant i \leqslant r} \sum_{1 \leqslant j \leqslant s} [\gamma_i \beta_j - \min(\gamma_i, \beta_j)],$$

where

$$I_q \otimes A^t = \operatorname{diag}(A^t, \dots, A^t) \in M_{lq},$$

$$B \otimes I_l = \operatorname{diag}(B_1 \otimes I_l, \dots, B_s \otimes I_l) \in M_{lq}, \quad B_j \otimes I_l \in M_{l\beta_j}.$$

Hence, $I_q \otimes A^t - B \otimes I_l$ is a block diagonal matrix, with blocks of sizes $\beta_1 l, \dots, \beta_s l$. Moreover, the j block has the form

$$C_j = \begin{pmatrix} A^t & & & \\ -I_l & A^t & & \\ & \ddots & \ddots & \\ & & -I_l & A^t \end{pmatrix} \in M_{l\beta_j}.$$

Then, $\operatorname{rank}(I_q \otimes A^t - B \otimes I_l) = \sum_{1 \leqslant j \leqslant s} \operatorname{rank} C_j$.

By permutations of rows and columns, the matrix C_j is equivalent to a matrix $\operatorname{diag}(D_{j1}, \dots, D_{jr})$ with

$$D_{ji} = \begin{pmatrix} A_i^t & & & \\ -I_{\gamma_i} & A_i^t & & \\ & \ddots & \ddots & \\ & & -I_{\gamma_i} & A_i^t \end{pmatrix} \in M_{\gamma_i \beta_j}.$$

Hence, $\operatorname{rank}(I_q \otimes A^t - B \otimes I_l) = \sum_{1\leqslant i\leqslant r}\sum_{1\leqslant j\leqslant s} \operatorname{rank} D_{ji}$, and from Lemma 4.9 we have the stated result. □

Next, we prove Theorem 4.8.

Proof of (i). From the structure of M we have that

$$\begin{aligned}\operatorname{rank} M \geqslant{}& \operatorname{rank}(I_d \otimes A^t - A \otimes I_d) + \operatorname{rank}(I_{n-d} \otimes B^t - B \otimes I_{n-d})\\ &+ \operatorname{rank}(I_{n-d} \otimes A^t - B \otimes I_d).\end{aligned}$$

It is obvious that condition (i) is verified when the matrix $C = 0$. So, from Proposition 4.7 and Lemma 4.10, condition (i) follows immediately.

Proof of (ii). Analogously, from the structure of M we have that

$$\begin{aligned}\operatorname{rank} M \leqslant{}& \operatorname{rank}(I_d \otimes A^t - A \otimes I_d) + \operatorname{rank}(I_{n-d} \otimes B^t - B \otimes I_{n-d})\\ &+ d(n-d).\end{aligned}$$

We begin studying the case where A and B have only one block, that is to say, $\gamma = d$ and $\beta = n - d$.

We will see that a matrix C with a 1 placed in the first row and last column verifies the condition of maximum rank.

By transformations of columns, we eliminate the last γ columns of $I_d \otimes A^t - A \otimes I_d$ and the last β columns of $I_{n-d} \otimes B^t - B \otimes I_{n-d}$. And by transformations of rows, we make the first γ rows of $I_d \otimes A^t - A \otimes I_d$ zero. Notice that the matrix $-I_\gamma$ of $-\overline{C} \otimes I_d$ is not modified with this operation.

Then, considering the linearly independent rows and by permutations of rows, we will have an upper triangular block matrix with identity matrices in the diagonal and rank equal to $\gamma^2 - \gamma + \beta^2 - \beta + \gamma\beta$. So, in this case, the theorem is proved.

It can be seen that, to prove the cases with more than one block in the matrices A and B, the same technique can be used considering a matrix $C = [C_{i,j}] \in M_{d,n-d}$ with $C_{ii} = N_{\gamma_i,\beta_i}$ if $i \leqslant \min(r, s)$ and $C_{i,j} = 0$ if $i \neq j$, where N_{γ_i,β_i} is the $\gamma_i \times \beta_i$ matrix with a 1 placed in the first row and last column. Therefore, the theorem holds in the general. □

Corollary 4.11. *For all* $(f, V) \in \mathscr{M}$, codim $\mathscr{O}(f, V) > 0$. *In particular, any pair of* $\mathscr{M}$ *is not structurally stable with regard to the considered equivalence relation.* (*We take here the definition of structurally stable given by Willems in* [9], *that is to say,* $x \in \mathscr{M}$ *is said to be structurally stable if* $\mathscr{M}/\sim$ *contains an open neighborhood of* x).

5. Endomorphisms having a fixed invariant subspace

In this section, we consider in the set of matrices of the form $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ the equivalence relation defined by the similarity action of the subgroup of Gl_n formed by matrices of the form $\begin{pmatrix} S_1 & S_2 \\ 0 & S_4 \end{pmatrix}$. Although the obtention of an explicit miniversal deformation in this context is an open problem, significant progress for its solution is made in [3]. Here, we give an alternative approach to this problem, based in Theorem 4.4.

This approach is based on the fact that the matrices of the form $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$, are those of M_n that keep invariant the subspace $\mathrm{Im}\begin{pmatrix} I_d \\ 0 \end{pmatrix}$. Let $V_0 = \mathrm{Im}\begin{pmatrix} I_d \\ 0 \end{pmatrix}$. We define $M_0 = \{f \mid (f, V_0) \in \mathscr{M}\}$, which is the vector space of dimension $n^2 - d(n-d)$ formed by the matrices of the form $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$, and we consider the injective map

$$
\begin{aligned}
&i : M_0 \to \mathscr{M} \\
&i(f) = (f, V_0).
\end{aligned}
$$

Given a smooth map $\varphi : \mathscr{U} \to \mathscr{M}$ with $\mathscr{U}$ an open set of $\mathbb{F}^k$, we denote $\varphi_i := \pi_i \circ \varphi$, $i = 1, 2$, where we recall that π_1 and π_2 are the natural projections of $\mathscr{M}$ on M_n and $\mathrm{Gr}_d(\mathbb{F}^n)$, respectively. We make use of the following lemma.

Lemma 5.1. *If $\mathscr{U}$ is a neighborhood of the origin of $\mathbb{F}^k$, for every smooth map $\varphi : \mathscr{U} \to \mathscr{M}$ with $\varphi(0) = (f_0, V_0)$, there exists a smooth map $\varphi^* : \mathscr{U} \to \mathrm{Gl}_n$ with $\varphi^*(0) = I_n$ such that*

$$\varphi^*(\lambda)[\varphi_2(\lambda)] = V_0$$

for all λ in an open neighborhood of the origin of $\mathscr{U}$.

Proof. In a neighborhood of the origin of $\mathscr{U}$, $\varphi_2(t) = \mathrm{Im}\begin{pmatrix} I_d \\ Q(\lambda) \end{pmatrix}$ with $Q(0) = 0$. Defining $\varphi^*(\lambda) = \begin{pmatrix} I_d & 0 \\ -Q_\lambda & I_{n-d} \end{pmatrix}$, the lemma holds. $\square$

Next proposition shows how a deformation in M_0 can be obtained through a deformation in $\mathscr{M}$.

Proposition 5.2. *Let $\varphi : \mathscr{U} \to \mathscr{M}$ be a versal deformation of (f_0, V_0) with regard to the action of Gl_n. Then, the map $\bar{\varphi} : \mathscr{U} \to M_0$, defined by $\bar{\varphi}(\lambda) = \varphi^*(\lambda) \circ \varphi_1(\lambda) \circ \varphi^*(\lambda)^{-1}$ is a versal deformation of f_0 with regard to the similarity action defined by the group $\{g \in \mathrm{Gl}_n \mid g(V_0) = V_0\}$.*

Proof. We first prove that $\bar{\varphi}(\lambda) \in M_0$ for all $\lambda \in \mathscr{U}$. In fact, from Lemma 5.1

$$\bar{\varphi}(\lambda)V_0 = (\varphi^*(\lambda) \circ \varphi_1(\lambda))(\varphi_2(\lambda)) \subset \varphi^*(\lambda)(\varphi_2(\lambda))$$

and, again from Lemma 5.1, we have that $\bar{\varphi}(\lambda)V_0 \subset V_0$ and so, $\bar{\varphi}(\lambda) \in M_0$.

Let $\bar{\psi} : \mathscr{V} \to M_0$ a smooth map with $\overline{\psi}(0) = f_0$.

For being φ versal, there exist $\mathscr{V}' \subset \mathscr{V}$ (neighbourhood of the origin), $\alpha : \mathscr{V}' \to \mathscr{U}$ with $\alpha(0) = 0$ and $\beta : \mathscr{V}' \to \mathrm{Gl}_n$ with $\beta(0) = I_n$, being α, β smooth maps, such that $\psi(\mu) = \beta(\mu) * \varphi(\alpha(\mu))$, and then,

$$(\bar{\psi}(\mu), V_0) = (\beta(\mu)\varphi_1(\alpha(\mu))\beta(\mu)^{-1}, \beta(\mu)[\varphi_2(\alpha(\mu))]).$$

Hence, according to the definition of φ^* we have that

$$\begin{aligned}\bar{\psi}(\mu) &= \beta(\mu) \circ \varphi^*(\alpha(\mu))^{-1} \circ \bar{\varphi}(\alpha(\mu)) \circ \varphi^*(\alpha(\mu)) \circ \beta(\mu)^{-1} \\ &= (\beta(\mu) \circ \varphi^*(\alpha(\mu))^{-1} * \bar{\varphi}(\alpha(\mu)).\end{aligned}$$

Defining

$$\varepsilon(\mu) := \beta(\mu) \circ \varphi^*(\alpha(\mu))^{-1},$$

we have that $\varepsilon(0) = \beta(0) \circ \varphi^*(0)^{-1} = I_n$, $\varepsilon(0)V_0 = V_0$ and one can check that

$$\bar{\psi}(\mu) = \varepsilon(\mu) * \bar{\varphi}(\alpha(\mu)).$$

So, according to Definition 4.1 we have that $\bar{\varphi}$ is versal. $\square$

We conclude that a versal deformation of $\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$ with regard to the similarity action of matrices of the same type, is given in terms of the versal deformation of the pair $\left(\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}, \mathrm{Im}\begin{pmatrix} I_d \\ 0 \end{pmatrix}\right)$ described in Theorem 4.4.

Acknowledgment

We are very grateful to F. Puerta for valuable comments and suggestions.

References

[1] V.I. Arnold, On matrices depending on parameters, Uspekhi Mat. Nauk. 26 (1971) 101–114.
[2] V.I. Arnold, Arnold's Problems, Springer, Berlin, 2004.
[3] A. Compta, J. Ferrer, F. Puerta, Miniversal deformations of marked matrices, Linear Algebra Appl. 361 (2003) 181–201.
[4] J. Ferrer, I. Garcia, F. Puerta, Brunovsky local form of a holomorphic family of pairs of matrices, Linear Algebra Appl. 253 (1997) 175–198.
[5] J. Ferrer, F. Puerta, Versal deformations of invariant subspaces, Linear Algebra Appl. 332–334 (2001) 569–582.

[6] I. Gohberg, P. Lancaster, L. Rodman, Invariant Subspaces of Matrices with Applications, Wiley, New York, 1986.
[7] U. Helmke, J.B. Moore, Optimization and Dynamical Systems, Springer-Publ., London, 1994.
[8] M.A. Shayman, On the variety of invariant subspaces of a finite-dimensional linear operator, Trans. AMS (274.2) (1982) 721–747.
[9] J.C. Willems, Topological classification and structural stability of linear systems, J. Differential Equations 35 (1980) 306–318.